DOCUMENT RESUME

ED 363 304 IR 016 423

AUTHOR Gibbs, William J.; Rattanapian, Vichuda
TITLE Pilot Testing Instructional Software with Video: Cross Cultural Implications.
PUB DATE 93
NOTE 8p.; In: Art, Science & Visual Literacy: Selected Readings from the Annual Conference of the International Visual Literacy Association (24th, Pittsburgh, PA, September 30-October 4, 1992); see IR 016 399.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Computer Assisted Instruction; Computer Software Development; *Computer Software Evaluation; Cultural Differences; *Evaluation Methods; Feasibility Studies; Foreign Countries; Foreign Students; Higher Education; Observation; *Pilot Projects; Social Bias; *Student Reaction; *Undergraduate Students; *Videotape Recordings
IDENTIFIERS Thailand; United States

ABSTRACT

To assist instructors with software development efforts, the Behavioral Learning Center at Pennsylvania State University's Smeal College of Business implemented a method to pilot test instructional software. The developed method was required to collect multichanneled information and was not limited to students' verbal responses to questions. Cultural diversity and personal bias also had to be eliminated. The adopted approach included video recording all pilot testing sessions, which allowed for a permanent visual record of the evaluation. The sample population consisted of five U.S. and five Thai undergraduate students at the university. The pilot testing videotape and rated each program on 4 semantic differential In the pilot study, the video technique allowed the observation that U.S. students tended to be more elaborate in their reactions, and generally more assertive than the Thai students. The capability to visualize the pilot testing sessions made the cultural differences much more observable. Two tables present student responses. (Contains 11 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 363 304

Pilot Testing Instructional Software with Video: Cross Cultural Implications

by
William J. Gibbs and Vichuda Rattanapian

Introduction

The challenges of ever expanding class sizes and associated class loads have forced many universities to investigate computer technology for education. In today's classroom it is commonplace to see computer and/or integrated multimedia workstation applications employed for instructional purposes in both presentational and self-instructional modes. More-over, there appears to be an interest on part of many instructors to be involved in developing such products. Few faculty, however, have the time to construct these systems. Given this problem, a remedy lies in authoring systems (Garrett, Noblitt, & Dominguez, 1989). The functionality and accessibility of many computer authoring systems has enabled some educators to assume the role of instructional software developer. Instructors can prepare class presentations which incorporate graphics, sound, animation and video. Similar self-instructional programs can be developed for students' independent use. As Suzuki (1987) notes, the novice author can develop Computer Assisted Instruction (CAI) in a relatively short period of time.

The ability to create fairly sophisticated software lessons using an authoring system has inspired individuals from an array of disciplines to develop their own instructional courseware. Unfortunately, what is sometimes missing from these lessons is the application of sound instructional design methodologies. Lessons, for example, are developed and implemented without some type of formative evaluation. In formative evaluation the program's worth can be determined and decisions about program revisions are made during its development (Gagne', Briggs, & Wager, 1992). These procedures enable one to "gain the capacity to confirm decisions pertaining to the program as valid, to improve efforts if required, or to modify thinking when necessary" (Hannafin & Peck, 1988, p. 299). In order to assist instructors with their software development efforts the Behavioral Learning Center at Penn State's Smeal College of Business Administration implemented a method to pilot test instructional software. According to Allessi and Trollip (1985, p. 384), "pilot testing is a process whereby representatives from a target audience use and test lessons while their progress and performance are monitored." The purpose of this article is to suggest some advantages to using video technology in pilot testing a CAI prototype. It examines the impact of visualizing subjects' operations in a lesson and the implications this has for improving instructional software, as well as pilot testing procedures. Moreover, the article discusses the impact of using video to assess the influence culture has on subjects' reactions to a CAI prototype and the pilot testing experience.

Background

The systematic design of instructional software increases its likelihood of successful implementation. All too often, however, trained instructional designers are not included on teams responsible for designing and developing programs (Gayeski, 1989). What can result is poorly designed instruction. What often appears intuitive to the untrained developer can mystify the student. Further, programs

1201623

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Alice D. Walker

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

that work flawlessly in a controlled setting can "crash" in the field. For these reasons, testing prior to courseware implementation is critical to its success. Courseware developed without being put through formative evaluation processes will likely create difficulty for students using it (Suzuki, 1987). This testing or evaluation is possibly the most frequently disregarded component of the systems model (Hannafin & Peck, 1988).

To ensure the adequacy of instructional software developed by instructors, the Behavioral Learning Center set up procedures for pilot testing programs. Three factors were considered when establishing an evaluation technique. First, the information collected needed to be multi-channeled and not limited to students' verbal responses to questions. Although verbal responses are worthwhile, there are other response types which can provide valuable information. Many channels of communication serve to exchange messages and are important such as facial expressions, eye contact, and body gestures (Wrightsman & Deaux, 1981). This is particularly important when observing international subjects. Language barriers can impede an observer's ability to interpret verbal responses. Therefore, nonverbal communication can aid the observer in understanding responses. "Nonverbal communication includes the space or distance between communications, touching, body messages, and voice characteristics" (Pederson, Wright, & Weitz, 1981, p. 247).

Second, because individuals of a particular culture will not perceive something totally alien to their cultural experience, or they will interpret it by some referent to their own cultural (Mangan, 1978), the target audience needed to be culturally diversified. This was important because many international students studying at Penn State are users of instructional software. Thus, the testing method had to be capable of documenting any effects on students' reaction patterns which were culturally based.

Third, the accounts of several observers monitoring the same event often vary due to personal biases (Van Dalen, 1979). For this reason, an instrument unaffected by personal bias was required.

The adopted approach included video recording all pilot testing sessions. Subjects' verbal, nonverbal and attitudinal reactions to the courseware were recorded, as well as their paths through the lesson. This technique provided the testing moderator with a permanent visual record of evaluation. The visual medium allowed the recording of multi-channeled communications and subsequent analysis of it by replaying the video tape.

Description of Population

The sample population used in pilot testing consisted of 5 U.S. students and 5 Thai students. The U.S. students were undergraduate and graduate part-time workers at the Smeal College of Business Administration at Penn State. The Thai students were all working on graduate degrees at Penn State and had been in the United States for at least one year but no longer than three years. All subjects were computer novices.

Procedure

Individually, subjects were given a brief introduction to the evaluation session and the program. They were told to think aloud as they used the courseware, to comment on program stimuli as it was presented and to ask questions when necessary. Subjects were left alone to run the program. The moderator monitored subjects' actions and took notes. If significant problems with any program aspect arose, then the counter time on the video tape recorder was marked so that point in the video could be accessed later for review. Subjects could terminate the program at any time. After each session, video tapes were reviewed.

General Observations

Six observations were made pertaining to this form of pilot testing. First, video recording provided a permanent visual and auditory record of the student-computer interaction. The video could be reviewed numerous times to identify subtle details about subjects, the software or the testing situation. Consequently, reviewers could identify problematic program areas more readily. For example, students' comments could be heard and their physical reactions seen along with what was presented on the computer screen. Such observations would be difficult to make with more traditional techniques.

Second, the ability to visualize subjects' reactions concurrently with program stimuli added a clearer understanding of what they liked, understood, disliked and/or found confusing about the courseware. As Allessi and Trollip (1985) point out, the amount of information relevant to students' understanding or confusion about a topic that can be obtained from observation is startling. It is the opinion of the researchers that in this particular case more detailed observations were possible because of the video taping.

Third, observer interruptions were fewer because the video provide a record of testing. The actions of subjects using a CAI lesson are often made rapidly. As a result, the observer may ask the individual to explain the action or to slow down. These interruptions can interfere with pilot testing. This was not the case in the present inquiry. Subjects could move through the courseware as they would under normal conditions. Moderators, for example, found it necessary to interrupt subjects only when they had completed a program segment without exploring all its options.

Fourth, since this method appeared to increase the amount of accessible data, more time was spent in analysis.

Fifth, video recording pilot testing sessions provided useful insights for improving pilot testing techniques. During testing, a moderator observed subjects from a separate room, and periodically asked them questions through an intercom system. This arrangement worked well for the U.S. subjects, but it proved problematic for the internationals. In the moderator's opinion some Thai subjects appeared uneasy with the testing situation. Consequently, the moderator moved into the testing room with these individuals. Possibly because of language differences the moderator prompted the Thai subjects more than the U.S subjects. Prompting was ineffective because it did not give the subject an opportunity to explore the program and its options independently of the moderator. In some cases it was difficult to determine if the subject's choice of a particular option was their own or the result of the moderator's prompting.

Sixth, video taping pilot testing sessions can be considerably more expensive to implement than simply observing students and writing down notes. It requires fairly sophisticated video equipment. The financial obligation of purchasing and maintaining such equipment for pilot testing can be prohibitive (Van Dalen, 1979).

Cultural Observations

People from varying cultural backgrounds will have both obvious and subtle differences which are not associated with ability (Woolfolk, 1990). Video recording evaluation sessions enabled reviewers to visualize how students from different cultures interacted with the courseware, the testing moderators and the testing approach. Consequently, the software and the evaluation approach could be modified to be more culturally sensitive.

A separate video tape was used to record each subject. All tapes were reviewed by three American evaluators and one Thai evaluator. The evaluators were told to write down any observable differences between the U.S. and Thai subjects pertaining to the following areas:

- Number of questions asked

- Type of questions asked
- Interaction with the moderator
- Interaction with the program
- Comfort with the testing environment
- Comfort with computer

The following list is a summary of observations made by the evaluators.

Thai subjects tended to:

1. Ask fewer questions than U.S. subjects. This data is shown in Table 1.

Table 1. Number of questions subjects asked during testing.

Number of Questions Asked		
U.S. Subjects		Thai Subjects
1.	27	15
2.	44	5
3.	26	15
4.	17	4
5.	15	9
Total	129	48

2. Hesitate before exploring program options more so than U.S. subjects. It appeared that the Thai subjects waited for a prompt from the moderator before moving on in the lesson. While this observation may be indicative of a cultural influence, the moderator's presence with some Thai subjects during testing may have affected their use of the program. This same effect may be present with the U.S. subjects given the presence of a moderator.

3. Present a more inhibited demeanor than the U.S. subjects. Cultural differences will result in communication differences. As Woolfolk (1990, p.122) indicates, for example, "some groups listen while giving strong affirmative nods and other nonverbal cues, while others listen without acknowledgement."

4. Comment less and make fewer suggestions than U.S. subjects. Thai subjects tended to wait for a question by the moderator before commenting on a program feature.

5. Follow the moderator's instructions verbatim. Their exploration of the courseware was limited to the moderator's directive. This observation may not be indicative of a cultural influence but an effect caused by the presence of the moderator. It should be noted, however, that even the Thai subjects who used the program without the moderator being present appeared to wait for a directive before proceeding.

6. Approach evaluation more seriously and more formally. Thai subjects tended to sit up straight in their chairs. Their body language appeared to suggest that they took this exercise seriousl.

7. Wait for assistance from the moderator when problems arose.

8. Maintain a great deal of patience when problems arose.

9. Question the utility and functionality of program features less often than the U.S. subjects.

10. Spend less time in testing than U.S. Subjects. This data is shown in Table 2.

Table 2. Amount of time each subject spent in testing.

Time in Testing (minutes)		
U.S. Subjects		Thai Subjects
1.	49.24	36.27
2.	48.43	24.30
3.	54.33	34.26
4.	32.45	17.47
5.	26.55	24.31
Average:	42.20	27.30

U.S. subjects tended to:

1. Ask numerous questions about the program and its features. Refer to Table 1.

2. Approach the evaluation task more assertively than Thai subjects. For example, when U.S. subjects encountered an option or program feature unfamiliar to them or one which they disliked, they generally commented immediately. U.S. subjects also appeared to express their reactions more directly and freely. For instance, they stated comments similar to the following: "I do not like the color of this table" or "The text in this screen is too small." or "This button doesn't work right".

3. Approach the courseware less conservatively than Thai subjects. During pilot testing U.S. subjects appeared to explore program options sooner than Thai subjects. Generally, they approached the software by clicking the mouse around the screen trying out various screen, buttons.

4. Present a more relaxed demeanor during pilot testing. While testing the prototype, the U.S. subjects leaned backwards in the chair, touched the computer screen and generally talked louder than Thai subjects.

There appeared to be no observable differences between the types of question asked by the Thai and U.S. subjects. The majority of questions focused on program interface issues. For example, subjects wanted explanations as to the meaning of various icons. They questioned the placement of symbols, tables, and textual information. Some subjects were more detailed than others in their observations, indicating very specific aspects of the courseware which needed improvement. However, such questions and observations occurred across both groups and seem to be based more on individual preferences than cultural differences.

The moderator's presence with some of the Thai subjects appeared to confound testing. It was observed that Thai subjects relied on and deferred to the moderator more so than the U.S. subjects. It is difficult to determine if this behavior resulted from cultural influences, language barriers or the moderator's presence. This is an important issue requiring further investigation. If international students are to use courseware, their input is vital in the CAI development stage. From the present study, there appear to be differences between international and U.S. subjects' pilot testing behavior which must be considered in order to collect useful data. The moderators in the current study sat with some of the internationals during testing to create a more comfortable environment and to help alleviate language difficulties. In attempting to accommodate these subjects the moderator may have curtailed subjects' desire to explore the courseware. Furthermore, after reviewing the video tapes several times it was clearly evident that some reactions reflected the personal biases of the moderator. From the moderator's perspective the Thai students directed their own paths through the courseware. They selected options independently. After reviewing the video tapes, however, it appeared that the moderator's questions directed the subjects. For example, the moderator would ask "what do you think of this option?" and based on that question a lesson path was taken. Without such a question the student may not have tested the option or would have chosen a different lesson path. This information is relevant for developers in understanding which stimuli are effective navigational cues. In retrospect, the moderator's support of the Thai students should have been restricted considerably. The researchers contend that this observation was possible as a result of visualizing pilot testing behavior via video tapes.

The most striking observation was that both groups differentiated themselves by the manner in which they reacted to the courseware. One needs to be cautioned that U.S. students and internationals may perceive a particular program feature in much the same way but display entirely different nonverbal and verbal responses.

Summary

Formative evaluation procedures are necessary for the effective implementation of any CAI program. This paper has suggested a method of pilot testing CAI prototypes using video technology. Video gives the researcher an unbiased instrument with which to visualize the student-computer interface. "Video recordings preserve details in a reproducible form, so that the full accounts of an incident may be studied repeatedly and intensively by the investigator and also may be checked by other research workers" (Van Dalen, 1979, p.165). While using video to observe subjects is not a new concept, its application in this study is unique. Subjects' reactions were recorded concurrently with what was presented on the computer screen.

Given that rapid assimilation of computer technology into the classroom and the accessibility of easy-to-use authoring systems, many educators are anxious to incorporate their software into class materials. This pilot testing technique enables educators to collect a variety of responses (verbal, physical, attitudinal) from subjects which can be used to gain greater insights into students' reactions towards courseware. The impact of being able to visualize subjects' reactions to software cannot be over-emphasized. It clearly added another dimension to evaluating the CAI prototype.

Since perception is influenced by culture (Mangan, 1978), a further consideration in testing software is how students from different cultures react to it. In the present study, U.S. subjects appeared more elaborate in their reactions and generally more assertive during testing than Thai subjects. While both groups were reactive, the manner in which they reacted differed. These observations are preliminary but seem to suggest that internationals interact and react differently during pilot testing than do U.S. subjects. Additional research is needed to investigate the numerous factors which can account for this. Nevertheless, it is apparent that the capability to visualize pilot testing sessions made cultural differences much more observable.

References:

Alessi, S.M. & Trollip, S.R. (1985). *Computer-Based Instruction: Methods and Development.* Prentice-Hall, Inc., New Jersey.

Gagne', R.M., Briggs, L.J., & Wager, W.W. (1992). *Principles of Instructional Design.* Harcourt Brace Jovanovich College Publishers, Orlando, FL.

Garrett, N., Noblitt, J., & Dominguez, F. (1989). Computers in Foreign Language Teaching and Research: A "New Humanism". *Computing Across the Curriculum: Academic Perspectives.* Academic Computing Publications, Inc., McKinney, Texas.

Gayeski, D.M. (1989). Why Information Technologies Fail.*Educational Technology*, 29(2), 9-16.

Hannafin, M.J. & Peck, K.L. (1988). *The Design, Development, and Evaluation of Instructional Software.* MacMillan Publishing Company, New York.

Mangan, J. (1978). Cultural Conventions of Pictorial Representation: Iconic Literacy and Education. *Educational Communications and Technology*, 26(3), 245-267.

Pederson, C. A., Wright, M.D. & Weitz, B.A. (1981). *Selling Principles and Methods.* Richard D. Irwin, Inc. Homewood, Illinois.

Suzuki, K. (1987). A Short-Cycle Approach to CAI Development: Three Stage Authoring for Practitioners.*Educational Technology*, 27(7), p19-24.

Van Dalen, D.B. (1979). *Understanding Educational Research*. McGraw-Hill Book Company, New York.

Wrightsman, Lawrence, S. & Deaux, K. (1981). *Social Psychology in the 80s*. Third Edition. Brooks/Cole Publishing Company Monterey, California.

Woolfolk, A.E. (1990). *Educational Psychology*. Prentice Hall, New Jersey.